

Bioinformatics Resources

Ewy A. Mathé

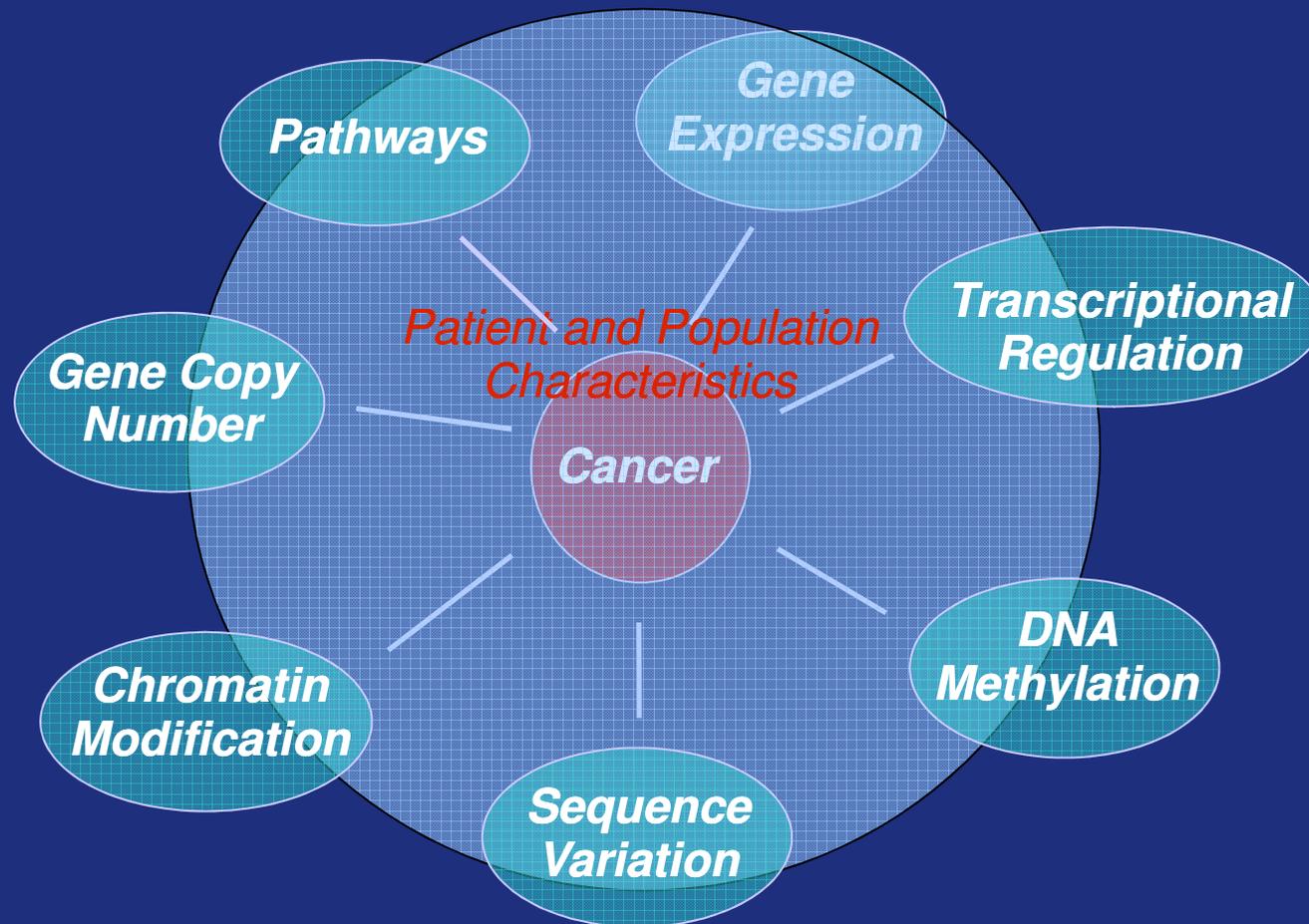
Laboratory of Human Carcinogenesis

NCI/NIH, Bethesda, MD

February 24th, 2011

Bioinformatics

Application and development of tools for querying large-scale datasets to answer biological questions.



Data



Overview

(1) Storage

(2) High-computing servers

(3) Software Available to CCR postdocs

(4) Training

(5) Bioinformatics Core

Storage

Memory has become cheap!

2 TB external hard drive ~ \$85



Cloud Storage:

- Online storage, data is stored on virtual servers
- Does not require maintenance or administration on the user end!
- Amazon, Google, Microsoft

Storage

Center for Information Technology Help Desk:

- Increased storage for data back-up
- Shared drive access

The screenshot shows a web browser window displaying the CIT Home page. The browser's address bar shows the URL <http://cit.nih.gov/>. The page header includes the CIT logo and the text "CENTER FOR INFORMATION TECHNOLOGY NATIONAL INSTITUTES OF HEALTH". A search bar is located in the top right corner. A navigation menu below the header contains links for "CIT Home", "CIT Service Catalog", "Information Security", "Support", "Science", "IT Policies", and "About CIT".

The main content area features three columns. The left column, titled "NIH IT SERVICE DESK", contains the following information:

- Need IT Support?**
[Submit a request online](#)
- Phone:**
301-496-4357 (local)
866-319-4357 (toll-free)
301-496-8294 (TTY)
- Web:**
ITServiceDesk.nih.gov

This section is circled in red. The middle column, titled "News & Alerts", includes links for "Hot News and Alerts from the NIH IT Service Desk" and "New Voice Mail System". The right column, titled "QUICK LINKS", lists various services such as "Application Hosting", "Antivirus", "CIT Link eNewsletter", "Co-Location Services", "Database Services", "NIH Electronic Forms", "Email", and "Encryption FAQ".

High-Computing Servers

NIH CIT Helix Systems Group (<http://helix.nih.gov>)

- High performance computing systems
- Training classes/seminars available
- \$26/month: account includes access to Biowulf for parallel processing (2200 compute nodes)



High-Computing Servers

Advanced Biomedical Computing Center, SAIC-Frederick (<http://isp.ncifcrf.gov/abcc/abcc-groups/>)

- Provides access to many publicly available resources (GenBank, PDB, SWISS-PROT, etc)
- Provides access to in-house software, ranging in applications from data mining to biomarker discovery
- Access is free of charge

Training: Intramural

NIH Library (nihlibrary.nih.gov, click on Bioinformatics)

- One on one consultation
- Training (online and classroom)
- Access to Bioinformatics tools and databases
- Access to Bioinformatics computers at the library

The screenshot shows a web browser window displaying the NIH Library Bioinformatics Program Staff page. The browser's address bar shows the URL: <http://nihlibrary.nih.gov/ResearchTools/Pages/biostaff.aspx>. The page features the NIH Library logo and navigation menus for Home, Library Services, Research Tools, Resource Training, and About Us. The main content area is titled "Bioinformatics Program Staff" and lists two staff members: Dr. Medha Bhagwat and Dr. Lynn Young. Dr. Bhagwat's bio mentions her PhD in biochemistry and her role in training at the NCBI Core-Bioinformatics Facility. Dr. Young's bio highlights her experience in computational biology and her work at the Center for Information Technology at NIH. A sidebar on the left provides quick links to various services like "Ask a Librarian", "Search Library Site", and "Quick Links". A right sidebar contains "Announcements" and "Upcoming Training Classes". The footer includes contact information for the NIH Library and logos for the Office of Research Services, National Institutes of Health, and U.S. Department of Health & Human Services.

Training: Intramural

CIT Training (<http://training.cit.nih.gov>):

Helix Training courses

Office of Science and Technology Partnerships

[https://ccrod.cancer.gov/confluence/display/CCR
OSTP/Home](https://ccrod.cancer.gov/confluence/display/CCR+OSTP/Home)

Supports access Bioinformatics software and
organizes training

NIAID Courses

Training: Extramural

- Seminar series at nearby universities
 - University of Maryland
(<http://www.cbcb.umd.edu/seminars/>)
 - George Mason University
(<http://binf.gmu.edu/colloq.html>)
- Cold Spring Harbor Laboratory Courses
- Sante Fe Institute courses:
<http://www.santafe.edu/events/workshops/index.php/>
- Google!

Bioinformatics Core

CCR Informatics Core (IFX)
<http://ccrifx.cancer.gov/>

- Provides Bioinformatics expertise and services to CCR investigators
- Requests submitted through the web, requires PI approval

The screenshot shows a web browser window displaying the CCR Informatics Core (IFX) website. The browser's address bar shows the URL <http://ccrifx.cancer.gov/apps/site/default>. The website header features the National Cancer Institute logo and the text "National Cancer Institute" and "U.S. National Institutes of Health | www.cancer.gov". The main content area includes the CCR IFX Informatics logo, a navigation menu with links for Home, Resources, FAQ, Staff, Subscribe, Projects, and Services, and a "New Project?" section. The "New Project?" section contains the text: "You may request a new project, or view your existing projects, by clicking on the [Projects](#) tab on the navigation bar." Below this is a "Latest Resources" section listing several tools and their dates: SAMtools (added 2011-01-11 10:00:09), Genomatrix Genome Analyzer (GGA) (added 2010-12-03 09:08:01), Galaxy (added 2010-12-03 09:06:13), CisGenome (added 2010-12-03 09:04:31), and PeakSeq (added 2010-12-03 09:00:53). To the right of the "Latest Resources" section is a "Site Search" section with a search input field and a magnifying glass icon. Below the search section is a "Latest FAQ" section listing several frequently asked questions and their dates: "What is the BAM file format?" (added 2011-01-11 09:38:14), "How do I submit a new request for analysis on my project?" (added 2010-12-28 15:42:47), "What steps are required for SNP calling?" (added 2010-12-03 10:44:50), "Where do I find mutation data for NCI-60 cancer cell lines?" (added 2010-12-03 09:03:46), and "Where can I find a guide for Microarray Data Analysis?" (added 2010-12-03 08:48:32).

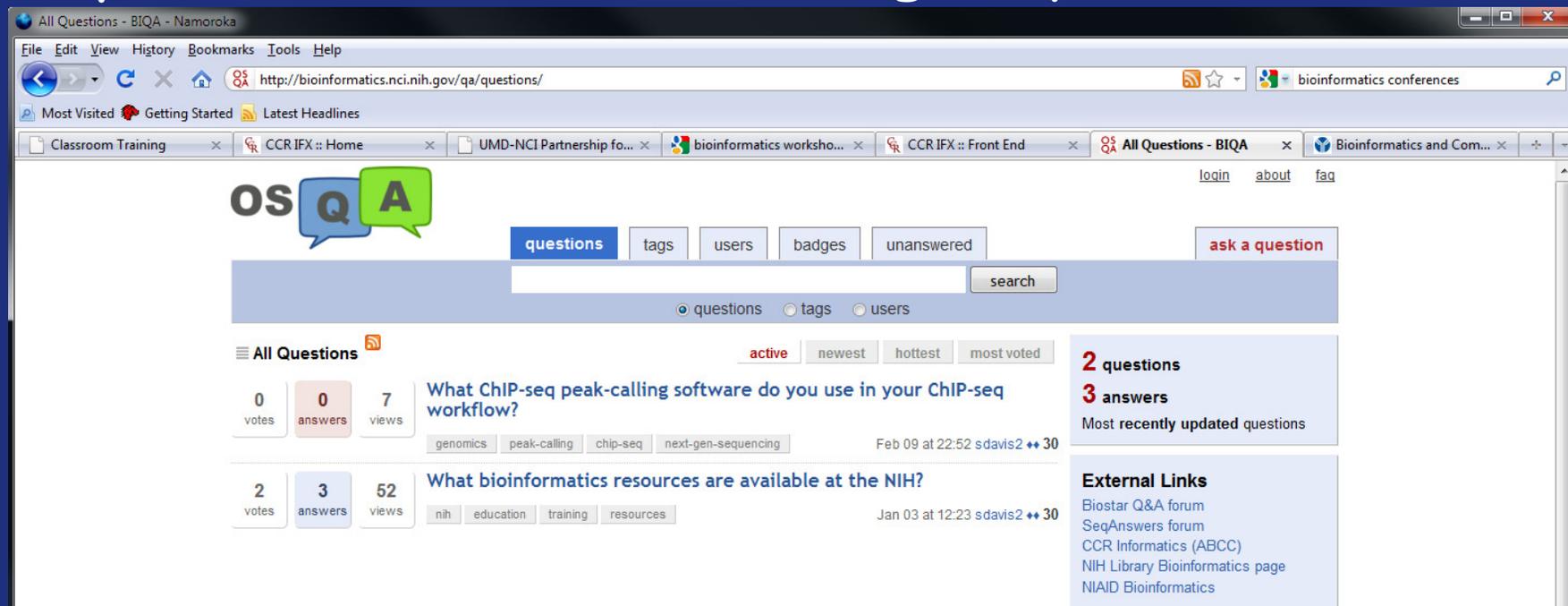
Other Resources

CCR Bioinformatics Wiki:

<https://ccrod.cancer.gov/confluence/display/CCRBioInfo/Home>

Bioinformatics QA Forum:

<http://bioinformatics.nci.nih.gov/qa>



The screenshot shows a web browser window displaying the Bioinformatics QA Forum. The browser's address bar shows the URL <http://bioinformatics.nci.nih.gov/qa/questions/>. The forum's logo, "OSQA", is visible at the top left. The page features a navigation menu with tabs for "questions", "tags", "users", "badges", and "unanswered", along with an "ask a question" button. A search bar is located below the navigation menu. The main content area displays a list of questions. The first question is "What CHIP-seq peak-calling software do you use in your CHIP-seq workflow?", posted by sdavis2 on Feb 09 at 22:52. It has 0 votes, 0 answers, and 7 views. The second question is "What bioinformatics resources are available at the NIH?", posted by sdavis2 on Jan 03 at 12:23. It has 2 votes, 3 answers, and 52 views. On the right side of the page, there is a sidebar with a "Most recently updated questions" section showing 2 questions and 3 answers, and an "External Links" section listing resources like Biostar Q&A forum, SeqAnswers forum, CCR Informatics (ABCC), NIH Library Bioinformatics page, and NIAID Bioinformatics.